

ABSTRACT

Self Tuning Database Retrieval Optimization Using Regression Functions

A system and method for accessing a relational database and estimating the selectivity of a query (e.g., an SQL query) in order to better predict the number of qualifying records for simple and complex queries. A dataset is created based upon queries
5 applied against the database by the user community. The dataset is populated with information related to query conditions and their respective combinations. A regression function reflecting correlations between query conditions is generated and used as a data mining model to calculate table-specific estimates for the cardinality of subsequent queries. An appropriate access method is then selected from a set of
10 available access methods based upon the number of estimated query-qualifying records. By periodically updating the regression model with FIFO managed queries, a self-tuning mechanism is achieved resulting in better selectivity/result size estimates for use in selecting access methods used in compiling subsequent SQL queries applied against the database.